

# Bayesian statistical methods for the analysis of multi-model climate predictions

**Doctoral Thesis****Author(s):**

Buser, Christoph Markus

**Publication date:**

2009

**Permanent link:**

<https://doi.org/10.3929/ethz-a-005927612>

**Rights / license:**

In Copyright - Non-Commercial Use Permitted

Diss. ETH No. 18448

# Bayesian Statistical Methods for the Analysis of Multi-model Climate Predictions

A dissertation submitted to  
ETH ZURICH

for the degree of  
Doctor of Sciences

presented by  
CHRISTOPH MARKUS BUSER  
Dipl. Math. ETH  
born Juli 03, 1978  
citizen of Känerkinden BL

accepted on the recommendation of  
Prof. Dr. Hansruedi Künsch, examiner  
Prof. Dr. Christoph Schär, co-examiner

2009

# Abstract

Current climate change projections are based on comprehensive multi-model ensembles of global and regional climate simulations. Application of this information to impact studies requires a combined probabilistic estimate taking into account the different models and their performance under current climatic conditions. Here, we present a Bayesian statistical model for the distribution of mean surface temperatures for control and scenario periods. The model combines observational data for the control period with the output of regional climate models (RCMs) driven by different global climate models (GCMs). The proposed Bayesian methodology considers both changes in mean temperature and interannual variability. In addition, unlike previous studies, our methodology explicitly considers model biases that are allowed to be time-dependent (i.e. to change between control and scenario period). More specifically, the model considers additive and multiplicative model biases for each RCM and introduces two plausible assumptions (“constant bias” and “constant relation”) about extrapolating the biases from the control to the scenario period. The resulting identifiability problem is resolved by using informative priors for the bias changes. A sensitivity analysis illustrates the role of the informative priors.

As an example, we present results for Alpine winter and summer temperatures for control (1961 - 1990) and scenario periods (2071 - 2100) under the SRES A2 greenhouse gas scenario. For winter, both bias assumptions yield a comparable mean warming of 3.5 to 3.6°C. For summer, the two different assumptions have a strong influence on the probabilistic prediction of mean warming, which amounts to 5.4°C and 3.4°C for the “constant bias” and “constant relation” assumptions, respectively. Analysis shows that the underlying reason for this large

uncertainty is due to the overestimation of summer interannual variability in all models considered. Our results show the necessity to consider potential bias changes when projecting climate under an emission scenario. These results are confirmed for other regions in Europe. Whether the “constant bias” and “constant relation” assumptions are different or similar depends not only on the season, but also on the region considered. Furthermore, we calculate predictions for the globe and the Northern Hemisphere using GCM output. The results are similar to those of the IPCC AR4 report.

In a first extension, the Bayes model is adapted to autocorrelated data. Serial correlations between neighbouring time periods appear if these periods are short, e.g. when analysing monthly mean values. The annual cycle is taken into account, both for the expectation values and for the development of the interannual variability for different seasons. There are some indications that such cycles are shifted or have different amplitudes under the scenario climate, i.e. that the annual maximum temperatures is shifted from July to August. Although these changes are smaller than the mean change, they are important for impact studies. There is no change of the correlation structure.

A second extension deals with spatially correlated data. Instead of spatial mean values we consider principal components of the correlated grid data. These components have a direct interpretation, i.e. they are the mean, the altitude of a station and a north-south and west-east gradient, respectively. The Bayesian analysis is done for each component separately. The resulting distributions can be used to predict a climate shift at a single location. These predictions are provided for five different stations in Switzerland for the scenario period. The corresponding prediction for the control period are in good agreement with the actual observed climatology for these stations. In addition to the individual analysis of each component, a multivariate model is introduced which allows one to account for correlations between the components.

Finally, we present an extension of the model to a situation with output from both, the RCM and the GCM simulations. This hierarchical Bayes model with random effects makes it possible to use RCMs that are driven by the same GCM run. It avoids problems with overconfidence due to the correlations of such RCM runs. The Bayes model estimates the different biases and assigns them to the GCMs or RCMs.

# Zusammenfassung

Prognosen für Klimaänderungen basieren auf verschiedenen Klimaprojektionen, die in einem Ensemble zusammengefasst und analysiert werden können, wobei die Güte der einzelnen Modelle berücksichtigt werden sollte. Für Anwender ist es oft wichtig, eine Verteilung einer Klimavariablen unter einem Szenario zu kennen. Wir stellen ein Bayes Modell vor, welches wir auf die Lufttemperatur in den Alpen anwenden. Das Modell kombiniert Informationen aus Beobachtungen mit regionalen Klimamodellen (RCMs), die von globalen Klimamodellen (GCMs) angetrieben werden. Solche Simulationen existieren jeweils für eine Kontrollperiode und eine Szenarioperiode. Neben der mittleren Klimatologie werden auch die Abweichungen von Jahr zu Jahr berücksichtigt. Im Gegensatz zu anderen Studien korrigieren wir die Modelle nicht nur für einen konstanten systematischen Fehler (Bias) in der Kontrollperiode, sondern schätzen eine mögliche Änderung, d.h. einen Bias, der über die Zeit variiert und von der jeweiligen Klimatologie abhängen kann. Das daraus entstehende Identifizierbarkeitsproblem wird mittels informativen Priors gelöst, wobei der Einfluss dieser Annahmen in einer Sensitivitätsanalyse untersucht wird. Dabei werden sowohl additive als auch multiplikative Fehler verwendet. Zwei Annahmen über die Extrapolation des Bias (“constant bias” und “constant relation”) in die Szenarioperiode werden verglichen.

Als Anwendung zeigen wir die Resultate für die Sommer und Winter Temperaturen im alpinen Raum für die Kontrollperiode (1961-1990) und die Szenarioperiode (2071-2100) für das SRES A2 Szenario. Während die beiden Extrapolationsmethoden im Winter ähnliche Resultate liefern, sind diese im Sommer stark unterschiedlich. Der Grund liegt in der systematischen Überschätzung der Variabilität im Sommer für alle Kli-

modellen. Wir zeigen daher, dass es wichtig ist, Annahmen über eine allfällige Entwicklung von systematischen Fehlern zu treffen und diese zu diskutieren. Die Annahme, dass Fehler konstant bleiben und vom jeweiligen Klimazustand unabhängig sind, darf nicht ohne genauere Untersuchung getroffen werden. Diese Resultate bestätigen sich auch in der Analyse anderer Regionen in Europa, wobei Unterschiede von Region zu Region und zwischen den Jahreszeiten eine Rolle spielen. Das Bayes Modell liefert auf globaler Skala eine Vorhersage, kombiniert aus verschiedenen GCMs, die vergleichbar mit den Resultaten des IPCC AR4 Reports vergleichbar ist.

Eine erste Erweiterung des Modells ermöglicht die Anwendung für autokorrelierte Daten, die auftreten wenn Mittelwerte von benachbarten, kurzen Zeitperioden untersucht werden, z.B. Monatsmittel. Die Erweiterung berücksichtigt den Jahresverlauf sowohl für das Mittel, aber auch für die Variabilität, die von Saison zu Saison unterschiedlich sein kann. Die Resultate zeigen eine mögliche Verschiebung des Jahreszyklus, auch wenn diese klein ist im Vergleich zu der Änderung des Mittelwertes. Weiter ist eine Intensivierung des Zyklus ersichtlich, d.h. Unterschiede zwischen Sommer und Winter werden noch verstärkt. Es ist keine Änderung der Korrelationsstruktur feststellbar.

Eine zweite Erweiterung ist notwendig, um nicht nur räumlich aggregierte Daten zu untersuchen. Dazu werden mittels Hauptkomponentenanalyse die wichtigen räumlichen Eigenschaften ermittelt. Diese sind der Mittelwert, die Höhe und ein Nord-Süd, bzw. West-Ost Trend. Für jede Grösse kann separat eine Prognose für Änderungen für die Szenarioperiode ermittelt werden und diese dann für eine gemeinsame Prognose eines einzelnen Gitterpunktes genutzt werden. Die Resultate von fünf verschiedenen Stationen in der Schweiz für die Kontrollperiode sind vielversprechend und entsprechen den tatsächlichen beobachteten Temperaturwerten. Neben der einzelnen Analyse der Hauptkomponenten wird ein multivariates Modell vorgestellt, welches die Korrelationen zwischen den verschiedenen Grössen berücksichtigt.

Schlussendlich zeigen wir eine Erweiterung des Modells auf Kombinationen von GCMs und RCMs, die es erlaubt, auch verschiedene RCMs zu verwenden, die vom selben GCM Lauf angetrieben werden. Diese hierarchische Erweiterung verwendet zufällige Effekte für die globalen Modelle, damit die Korrelationen zwischen den regionalen Modellen, die vom selben GCM angetrieben sind, berücksichtigt werden.